

# Shuffling

---

Nipun Batra

IIT Gandhinagar

August 3, 2025

# Consider this dataset

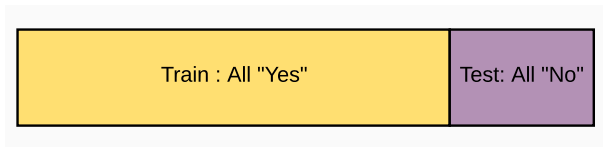
First 80 examples are of class "Yes"

Remaining 20 examples are of class "No".

<b>Serial Number</b>	<b>...</b>	<b>Class</b>
1		Yes
2		Yes
3		Yes
.		.
.		.
80		Yes
81		No
.		.
.		.
100		No

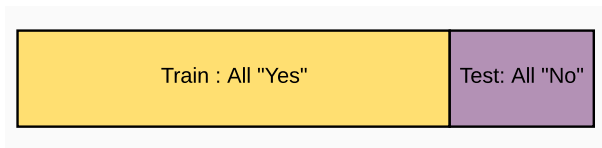
## Consider this dataset

While using an 80-20 train-test split, we will get the distribution shown below



# Consider this dataset

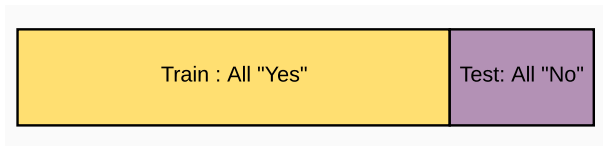
While using an 80-20 train-test split, we will get the distribution shown below



Will we learn anything useful in this scenario?

## Consider this dataset

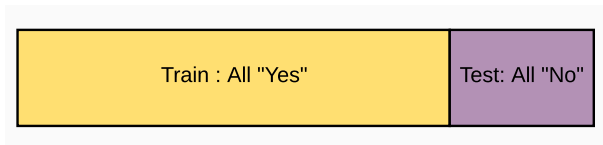
While using an 80-20 train-test split, we will get the distribution shown below



Will we learn anything useful in this scenario? No :(

## Consider this dataset

While using an 80-20 train-test split, we will get the distribution shown below



Will we learn anything useful in this scenario? No :(

**Solution:** Shuffle before learning

# Why shuffle for SGD?

We can fall into a loop!

SGD on point 1 :  $\theta_0 + 0.2, \theta_1 - 0.2$

SGD on point 2 :  $\theta_0 - 0.2, \theta_1 + 0.2$

Biased learning as point 2 follows point 1.